

Published in final edited form as:

J Mol Biol. 2007 June 1; 369(2): 334–342.

Molecular Implications of Evolutionary Differences in CHD Double Chromodomains

John F. Flanagan, Bartlomiej J. Blus, Daesung Kim, Katrina L. Clines, Fraydoon Rastinejad, and Sepideh Khorasanizadeh

Department of Biochemistry and Molecular Genetics, University of Virginia Health System, Charlottesville, VA 22908, USA.

Abstract

Double chromodomains occur in CHD proteins which are ATP-dependent chromatin remodeling factors implicated in RNA polymerase II transcription regulation. Biochemical studies suggest important differences in the histone H3 tail binding of different CHD chromodomains. In human and *Drosophila*, CHD1 double chromodomains bind lysine 4-methylated histone H3 tail which is a hallmark of transcriptionally active chromatin in all eukaryotes. Here, we present the crystal structure of the yeast CHD1 double chromodomains, and pinpoint its differences with that of the human CHD1 double chromodomains. The most conserved residues in these double chromodomains are the two chromoboxes that orient adjacently. Only a subset of CHD chromoboxes can form an aromatic cage for methyllysine binding, and methyllysine binding requires correctly oriented inserts. These factors preclude yeast CHD1 double chromodomains to interact with the histone H3 tail. Despite great sequence similarity between the human CHD1 and CHD2 chromodomains, variation within an insert likely prevents CHD2 double chromodomains to bind lysine 4-methylated histone H3 tail as efficiently as in CHD1. By using the available structural and biochemical data we highlight the evolutionary specialization of CHD double chromodomains, and provide insights about their targeting capacities.

Keywords

Double chromodomains; CHD; histone H3; lysine methylation; transcription; epigenetic signaling

Introduction

CHD (chromo-ATPase/helicase-DNA-binding) proteins are uniquely structured to encode double chromodomains, and these domains always occur N-terminal to a conserved SNF2 domain (Figure 1) ¹. A recent review discusses how multiple subfamilies of SNF2-containing proteins, found in eukaryotes, may perform ATP-dependent protein translocation on DNA, and promote chromatin remodeling ². *In vitro* studies have shown purified CHD1 proteins from *Saccharomyces* and *Drosophila* to exhibit ATP-dependent nucleosome repositioning to organize oligonucleosomes in a structure that is compatible with gene expression during transcription ^{3; 4}. All SNF2 containing proteins and their complexes contain other conserved domains that are expected to influence their repertoire of chromatin remodeling functions by

Corresponding author: Sepideh Khorasanizadeh E-mail: khorasan@virginia.edu Phone: (434) 243-4817 Fax: (434) 924-5069.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Accession Number: The coordinates have been deposited in the Protein Data Bank under ID code 2H1E.

forming specific protein-protein and protein-DNA contacts. In vertebrates, nine CHD variants contribute to the regulation of transcription, whereas *Saccharomyces* yeast encodes one CHD protein. Yet, molecular differences between distinct CHD proteins and their chromatin targeting potentials are poorly understood⁵. It is necessary to understand the structure and sequence differences among CHD proteins to establish a better understanding of their interaction and chromatin remodeling potentials.

A simplistic classification of CHD proteins based on their conserved domain organization suggests these proteins can be clustered into three classes as shown in Figure 1. For example, all CHD genes from unicellular organisms are similar to human CHD1, and distinct from two other classes that are encoded in multi-cellular organisms; these are represented by human CHD4 and human CHD7 genes in Figure 1. Human CHD3 and CHD4 integrate within multi-subunit complexes termed NuRD (nucleosome remodeling histone deacetylase; for review see; 6; 7; 8), and human CHD7 regulates differentiation and development of several organs^{9; 10}. Genetic studies have shown that *Saccharomyces* CHD1 is present at transcribed genes, and its chromodomains are essential for this function¹¹. Although regions of transcribed genes in all eukaryotes are rich in lysine 4 methylation on the histone H3 tail (H3K4me), neither the chromodomains nor the intact *Saccharomyces* CHD1 protein interacts with the H3K4me peptide^{12; 13; 14; 15}.

This is in stark contrast with the specific mode of interaction we previously identified and characterized for binding of the human CHD1 to the H3K4me peptide. To better understand the molecular features of the CHD1 double chromodomains, we have determined the atomic structure of this region in *Saccharomyces* CHD1 to delineate its differences with the human CHD1 structure. We have also generated a phylogenetic tree of CHD proteins based on their chromodomain segments, which suggests distinct functional specialization during evolution. We show 21-residue chromoboxes form the conserved core in each chromodomain of all CHD proteins, and propose diversity is generated by distinct insert regions outside the chromobox sequences. As the targets of the majority of CHD double chromodomains, including *Saccharomyces* CHD1, remain unidentified, our analysis offers a practical guide to further investigate the targeting and contribution of the chromodomain region to the function of the CHD proteins.

Results and Discussion

Differences between human and yeast CHD1

To explain the observed functional differences between human and yeast CHD1 chromodomains, we solved the crystal structure of the *Saccharomyces* CHD1 double chromodomains at 2.2 Å resolution (Table 1). Figure 2 shows this structure in comparison with that of the human CHD1 in complex with the methylated histone H3 tail. In both structures, each set of the secondary structure elements corresponding to a chromodomain assembles similar to the prototypical HP1 chromodomain¹³. However, sequence insertions in between the conserved regions in chromodomain 1 lead to a substantially larger chromodomain. A universal insert in CHD chromodomain 1 is insert 2, which serves to block peptide binding as seen in HP1 and Polycomb chromodomains¹³. The CHD double chromodomains arrange perpendicular to each other using a helical linker. Interestingly, the inter-chromodomain linker segments of *Saccharomyces* CHD1 and human CHD1 exhibit only 18% sequence identity, but both use 33 residues to fold similarly (Figure 2C). The structure of *Saccharomyces* CHD1 suggests chromodomains 1 and 2 may also cooperate to bind a partner.

Human and *Saccharomyces* CHD1 exhibit a major difference in the folding of their insert 1 region within chromodomain 1 (Figure 2). In *Saccharomyces*, insert 1 folds adjacent to the $\alpha 2$ region. Interestingly, two unconserved cysteine residues in insert 1 and $\alpha 2$ face each other

at a distance compatible with a disulfide bond, and this could further stabilize the close positioning of insert 1 to $\alpha 2$ only in *Saccharomyces* CHD1. These features in chromodomain 1 of *Saccharomyces* CHD1 have also been noted in the recently reported solution structure of this domain¹⁵. Unlike in *Saccharomyces*, human insert 1 forms the inter-chromodomain surface, and directly contributes to the histone H3 tail binding¹³. Additionally, in human CHD1, two tryptophans (322 and 325) in chromodomain 1 form an aromatic cage around methyllysine 4 of the H3 tail (Figure 2B)¹³. In *Saccharomyces* CHD1, there is a substitution of the second tryptophan with a Glu. To determine if Glu replacement with a Trp would impart methyllysine binding to *Saccharomyces* CHD1, we performed mutagenesis in conjunction with fluorescence anisotropy binding studies.

Figure 3c shows that the mutant *Saccharomyces* CHD1 E220W does not interact with the lysine 4 methylated histone H3 tail, suggesting that differences in other regions, particularly the folding of insert 1 dramatically affect binding capacities of CHD1 proteins. To date, a specific binding partner for *Saccharomyces* CHD1 double chromodomains has not been reported, but hints about such partners are revealed from interactome studies. *Saccharomyces* CHD1 has been identified in complexes with casein kinase II as well as a group of proteins that facilitate elongation for transcription by RNA polymerase II^{16; 17}. It is likely that the double chromodomains of *Saccharomyces* cooperate to accommodate binding of a protein partner at the juncture of chromodomains.

Structure of the Chromobox Homology Region

An alignment of the sequences of all CHD variants with non-CHD chromodomains reveals their most conserved region is limited to 21 residues that forms a conserved core in each chromodomain (Figure 3A,B and supplementary figure S1). We reintroduce the term chromobox to refer to this homology region. Originally the term chromobox was used to imply the nucleotide sequence that encodes 37 homologous amino acids in the chromodomains of HP1 and Polycomb genes¹⁸. The 21-residue chromobox in CHD1 proteins is highly related in structure to that region in HP1 and Polycomb proteins¹³. When aromatic residues are present at both positions 5 and 8 of the chromobox then formation of a two-residue aromatic cage for binding a methyllysine is predicted (Figures 2· 3b, and supplementary figure S1). We note that human CHD1 chromobox 1 meets this requirement, whereas neither human CHD1 chromobox 2 nor *Saccharomyces* CHD1 chromoboxes meets this requirement. The CHD chromoboxes form highly related structures due to significant sequence identity; there is 33% sequence identity between chromoboxes 1 and 2 in *Saccharomyces* CHD1. Also, the human and *Saccharomyces* CHD1 proteins have 62% identity in chromobox 1 and 43% sequence identity in chromobox 2 regions. Figure 3B shows that the absolutely conserved Tyr 10 in chromobox 2 could not cooperate with Trp 5 to form a novel aromatic cage as was predicted in a study that showed chromodomain 2 in *Saccharomyces* CHD1 binds lysine-4 methylated H3 tail¹⁹. Therefore, *Saccharomyces* CHD1, and highly likely, any other CHD protein could not use the chromodomain 2 region to assemble an aromatic cage for interaction with a lysine-methylated peptide.

Phylogenetic Classification of CHD Proteins According to Chromodomains

To understand the relationship among the diverse CHD proteins, we prepared a phylogenetic tree on the basis of their aligned chromodomain sequences (Figure 4). A total of 34 CHD proteins were included in our sequence alignment. We observed distinct patterns of similarities that suggest the presence of major evolutionary divisions in CHD proteins. This led to a clear split among CHD1 proteins where we found those from unicellular organisms to form a separate group (class B) from the CHD1 proteins of multi-cellular organisms (class A). Therefore, H3K4me binding is widely used in class A. Interestingly, in vertebrates, a CHD2 gene also belongs to class A that exhibits high sequence identity with human CHD1 (Figures

2C and 4). We noted a difference in the insert 3 region in chromodomain 2, as this region is 15 residues longer in CHD2 and likely interferes with peptide binding surface. To determine whether changes in insert 3 influence target selectivity, we prepared the double chromodomain fragment of human CHD2, and performed fluorescence anisotropy binding assays with methylated histone peptides. Figure 3c shows human CHD2 interacts with the H3K4me peptide 30-fold weaker than that detected for human CHD1. Furthermore, the presence of lysine 9 acetylation, arginine 2 methylation or lysine 36 methylation does not improve the binding affinity of human CHD2 for H3K4me peptide. While the biochemical function of CHD2 is unknown, mouse CHD2 gene is shown to have a role in development and survival²⁰. Therefore, we suggest an important role for insert 3 extension is to regulate H3K4me binding by CHD2 protein in a manner coordinated with complex stages of transcription. One possible mechanism may involve the action of a temporally regulated kinase that phosphorylates CHD2 to relieve insert 3 interference with histone tail binding.

The organisms listed in class B encode one CHD protein, except in *Schizosaccharomyces* where there are two CHD proteins called HRP1 and HRP3 (Figure 4). The double chromodomains of HRP1 and HRP3 are highly related to each other, yet HRP3 chromodomain 1 has aromatic residues at both positions 5 and 8 of the chromobox. Interestingly, the insert 1 region in both HRP proteins is shorter than that in *Saccharomyces* CHD1, suggesting although HRP3 may bind a lysine-methylated peptide the contribution of insert 1 to binding may be distinct from that used by human CHD1 (Figure 1B, and supplementary Figure S1).

Distinct from the two CHD1 classes is class C CHDs that includes CHD3 through CHD5, sometimes referred to as Mi-2 proteins (Figures 4 and supplementary Figure S1). These CHDs typically reside in multisubunit NuRD complexes, and contribute to histone deacetylation (for review see;⁶). Chromodomain 1 in this class always contains a hydrophobic residue at position 8 of the chromobox, precluding a methyllysine binding function via an aromatic cage. The inter-chromodomain linker is 10 residues longer than that in CHD1 proteins (supplementary Figure S1), suggesting an altered organization of the tandem chromodomains. The chromodomains of *Drosophila* Mi-2 are shown not to interact with methylated H3 tail, and instead are implicated in DNA binding that is required for nucleosome recognition²¹. The structure or affinity of such DNA binding is not yet determined, and DNA binding does not appear to be the function of CHD1 chromodomains, as we did not detect it for *Saccharomyces* CHD1 and human CHD1 (data not shown). Furthermore, in contrast to the inter-chromodomain linker in CHD1, class C linkers contain substantially more lysines and arginine residues, indicating a potential for using a unique surface of interaction for DNA binding. Interestingly, majority of class C CHDs contain two PHD fingers that always occur N-terminal to the double chromodomains (Figure 1). The PHD finger has also been implicated in lysine-methylated histone tail binding involving an aromatic cage^{22; 23; 24}. A CHD3 PHD finger was recently shown to bind H3K36me peptide using a pull down assay²³. It remains to be shown whether PHD fingers and class C chromodomains can cooperate to form a unique interaction with nucleosomes.

Another subfamily is class D that includes CHD7 (Figures 1 and 4), which is implicated in the CHARGE syndrome. CHARGE syndrome is a common cause of congenital anomalies affecting several tissues in humans. Mutations in the CHD7 gene of individuals with the CHARGE syndrome are believed to account for the disease²⁵. More recently the phenotypic spectrum of human mutations along CHD7 gene, including those in the chromodomain region, has been reported for patients with CHARGE syndrome²⁶. Other members of class D also appear to contribute to tissue and developmentally specific chromatin regulation. For example, CHD8 (or Duplin) is associated with beta catenin-mediated gene expression^{27; 28}, whereas CHD9 is associated with ligand-dependent transcription by nuclear receptors²⁹. Class D CHDs contains chromodomains with different aromatic residues at positions 5 and 8 of

chromobox 1 (Tyr at position 5 and a Phe/Tyr at position 8; see Figures 3a). They also exhibit very long insert 1 regions, suggesting class D may form an aromatic cage to bind methyllysine at the juncture of the two chromodomains (supplementary Figure S1A). Interestingly, this class exhibits an inter-chromodomain linker that is 13 residues shorter than the one in A and B classes (supplementary Figure S1B). Additional studies are necessary to determine whether reducing the linker impacts the tandem chromodomain arrangement and lead to new surfaces of inter-molecular interaction.

Finally, found four CHD proteins that exhibit dramatic divergence in their double chromodomain sequences. As such their evolutionary order of appearance could not be judged by our analysis. These are represented by two uncharacterized members from Arabidopsis, one from *Giardia lamblia* and one from *Plasmodium falciparum* (Figure 4). *Giardia lamblia* causes gastroenteritis in mammals, manifesting itself with severe diarrhea and abdominal cramps in humans. *Plasmodium falciparum* causes malaria, which is considered a major threat to human populations. The analysis of single nucleotide polymorphism for *Plasmodium falciparum* has previously revealed the divergence time for this organism to coincide with the start of human population expansion and consistent with a genetically complex organism able to evade host immunity³⁰. Despite dramatic sequence differences in the CHDs of *Plasmodium* and *Giardia*, their two chromoboxes are highly conserved as shown in Figure 3a. Moreover, the organization of conserved domains in both of these CHD proteins resembles that of CHD1 (Figure 1), suggesting they have a fundamental role in transcription regulation. As with *Saccharomyces* CHD1, *Plasmodium* and *Giardia* CHD proteins do not have the necessary combination of aromatic residues to bind a lysine-methylated peptide (Figure 3A).

Conclusion

In vitro studies have demonstrated that *Drosophila* CHD1 can act alone to organize nucleosome arrays in a form that is compatible with chromatin during transcription (i.e., a larger inter-nucleosome spacing and lack of incorporation of the linker histone H1)⁴. During transcription by RNA polymerase II, nucleosomes do not fall apart, but instead become displaced in a carefully regulated manner to comply with initiation, elongation and termination stages (for review see;³¹). We suggest CHD proteins manage nucleosome repositioning at various stages of transcription by responding to important epigenetic queues. We suspect the double chromodomains of CHD proteins assist this process and help to organize distinct CHD proteins at various sites of gene activity. While some unicellular organisms have only one class of CHD proteins, in worms, flies, plants and vertebrates, three distinct classes exist and all are implicated in transcription regulation. For class A CHDs we indicated their localization involves binding of their double chromodomains to the lysine 4 methylated H3 tail (for review see;³²). This modification typically correlates with transcription start sites as well as throughout the coding regions. Class C CHDs may use PHD fingers to recognize the lysine 36 methylated H3 tail, and use double chromodomains for DNA binding^{21; 23}. The H3K36me modification is also associated with transcription. As class C CHDs are components of histone deacetylating complexes, they may be related to the Rpd3S histone deacetylating complex that has been shown to localize at transcription termination sites in *Saccharomyces*. Within Rpd3S, a chromodomain protein called Eaf3 recognizes the H3K36me modification^{33; 34}.

Class D CHDs have double chromodomains that appear to recognize a protein partner bearing a methyllysine. In *Drosophila*, both CHD1 and Kismet-L facilitate transcription by RNA polymerase II^{35; 36}. Localizations of CHD1 and Kismet-L are distinct during transcription stages, and Kismet-L populates transcription sites prior to CHD1 arrival³⁶. Mammalian CHD6 is also shown to co-localize with RNA polymerase II during both preinitiation and elongation stages of transcription³⁷. Additional studies are required to understand the molecular mechanism(s) that various CHD proteins exert to prepare chromatin for RNA polymerase II

action. This is likely to be very important for the understanding of the basic transcription apparatus as well as the more complex tissue and cell specific gene transcription in multi-cellular organisms.

The appropriate mode of CHD localization to transcription sites appears to have dramatic consequences as mutations in CHD7 protein, that are implicated in the CHARGE syndrome, often lead to premature stop codons²⁶. Furthermore, malaria and *Giardia* parasites encode substantially different CHD proteins. For *Plasmodium falciparum* the complete genome is characterized and annotated³⁸. Moreover, the protein interaction network within malaria has been extensively studied, and resulted in identifying potential partners for the malaria CHD protein³⁹. Interestingly, a recent report indicated that dual infection with HIV and malaria fuels the spread of both diseases, and these diseases together cause the deaths of over 4 million people per year in Africa⁴⁰. The chromatin machinery of malaria likely contributes to this factor as malaria, which is not sexually transmitted, has dramatically impacted the process of spreading HIV infection.

Materials and Methods

Cloning and protein purification

All expression constructs contain an N-terminal 6xHis-tag, were cloned into the BamHI/NdeI sites of the pET11a vector, expressed in BL21(DE3) *E. coli* (Novagen), and purified by Ni²⁺-affinity chromatography (Qiagen). The double chromodomain constructs express residues 174-339 (for *Saccharomyces* CHD1) and residues 257-447 (for human CHD2) of the corresponding CHD genes. The E220W point mutation in *Saccharomyces* CHD1 was prepared using the QuikChange mutagenesis kit (Stratagene).

Binding assays

For fluorescence polarization, 100 nM of fluorescein-labelled peptide (prepared and used as previously described⁴¹) was used in 50 mM sodium phosphate pH 8.0, 25 mM NaCl, 5 mM TCEP (the binding assays involving human CHD2 contained an additional 25 mM NaCl). The peptides used in binding assays were H3K4me3 [ARTK(me3)QTARKSTGGKAY], H3K36me3 [APSTGGVK(Me3)KPHRY], H3K4me3K9acK14acK18ac [ARTK(me3)QTARK(ac)STGGK(ac)APRK(ac)QLAY], H3K4me3R2me2a [AR(me2a)TK(me3)QTARKSTGGKAY], and H3K4me3K36me3 [ARTK(me3)QTARKSTGGKAPRKQLATKAARKSAPATGGVK(me3)KPHRY].

Crystallization, data collection and Structure determination—Protein crystals were grown by hanging drop vapor diffusion at 10 °C with 6 mg/ml protein in 10 mM Bis-Tris Propane (BTP) at pH 8.0, 12 mM NaCl, 5 mM TCEP, 0.9 M ammonium sulfate and 3.5% isopropanol. The mother liquor was 1.8 M ammonium sulfate and 7% isopropanol, and the crystals were cryoprotected in 5 % isopropanol, 1 M ammonium sulfate and 35% ethylene glycol. These crystals diffracted to 2.2 Å in space group I41. Native diffraction data were collected at the APS SER-CAT 22-ID beamline and indexed with HKL2000⁴². The structure was phased using MOLREP⁴³ and using the human CHD1 tandem chromodomains as the search model. The structure was built with ArpWarp⁴⁴ followed by manual model building in Coot⁴⁵. The final structure was refined using RefMac5 restrained TLS refinement⁴⁶. All ribbon diagrams were drawn in PyMol⁴⁷, and the structure superpositions were generated by Coot.

Construction of a CHD phylogenetic tree

We selected 34 unique CHD proteins in the protein sequence database. The regions of double chromodomains were then aligned using CLUSTALW⁴⁸. This chromodomain alignment was then used to calculate the tree by maximum likelihood method in the PHYLIP package⁴⁹.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We thank William R. Pearson for advice on phylogenetic analysis. This work was supported by NIH grant GM070558 to S.K..

References

1. Woodage T, Basrai MA, Baxevanis AD, Hieter P, Collins FS. Characterization of the CHD family of proteins. *Proc Natl Acad Sci U S A* 1997;94:11472–7. [PubMed: 9326634]
2. Flaus A, Martin DM, Barton GJ, Owen-Hughes T. Identification of multiple distinct Snf2 subfamilies with conserved structural motifs. *Nucleic Acids Res* 2006;34:2887–905. [PubMed: 16738128]
3. Stockdale C, Flaus A, Ferreira H, Owen-Hughes T. Analysis of nucleosome repositioning by yeast ISWI and Chd1 chromatin remodeling complexes. *J Biol Chem* 2006;281:16279–88. [PubMed: 16606615]
4. Lusser A, Urwin DL, Kadonaga JT. Distinct activities of CHD1 and ACF in ATP-dependent chromatin assembly. *Nat Struct Mol Biol* 2005;12:160–6. [PubMed: 15643425]
5. de la Serna IL, Ohkawa Y, Imbalzano AN. Chromatin remodelling in mammalian differentiation: lessons from ATP-dependent remodellers. *Nat Rev Genet* 2006;7:461–73. [PubMed: 16708073]
6. Bowen NJ, Fujita N, Kajita M, Wade PA. Mi-2/NuRD: multiple complexes for many purposes. *Biochim Biophys Acta* 2004;1677:52–7. [PubMed: 15020045]
7. Becker PB, Horz W. ATP-dependent nucleosome remodeling. *Annu Rev Biochem* 2002;71:247–73. [PubMed: 12045097]
8. Feng Q, Zhang Y. The NuRD complex: linking histone modification to nucleosome remodeling. *Curr Top Microbiol Immunol* 2003;274:269–90. [PubMed: 12596911]
9. Brunner HG, van Bokhoven H. Genetic players in esophageal atresia and tracheoesophageal fistula. *Curr Opin Genet Dev* 2005;15:341–7. [PubMed: 15917211]
10. Shaw-Smith C. Oesophageal atresia, tracheo-oesophageal fistula, and the VACTERL association: review of genetics and epidemiology. *J Med Genet* 2006;43:545–54. [PubMed: 16299066]
11. Simic R, Lindstrom DL, Tran HG, Roinick KL, Costa PJ, Johnson AD, Hartzog GA, Arndt KM. Chromatin remodeling protein Chd1 interacts with transcription elongation factors and localizes to transcribed genes. *EMBO J* 2003;22:1846–56. [PubMed: 12682017]
12. Santos-Rosa H, Schneider R, Bernstein BE, Karabetsou N, Morillon A, Weise C, Schreiber SL, Mellor J, Kouzarides T. Methylation of histone H3 K4 mediates association of the Isw1p ATPase with chromatin. *Mol Cell* 2003;12:1325–32. [PubMed: 14636589]
13. Flanagan JF, Mi LZ, Chruszcz M, Cymborowski M, Clines KL, Kim Y, Minor W, Rastinejad F, Khorasanizadeh S. Double chromodomains cooperate to recognize the methylated histone H3 tail. *Nature* 2005;438:1181–5. [PubMed: 16372014]
14. Sims RJ 3rd, Chen CF, Santos-Rosa H, Kouzarides T, Patel SS, Reinberg D. Human but not yeast CHD1 binds directly and selectively to histone H3 methylated at lysine 4 via its tandem chromodomains. *J Biol Chem* 2005;280:41789–92. [PubMed: 16263726]
15. Okuda M, Horikoshi M, Nishimura Y. Structural Polymorphism of Chromodomains in Chd1. *J Mol Biol* 2007;365:1047. [PubMed: 17098252]
16. Gavin AC, Aloy P, Grandi P, Krause R, Boesche M, Marzioch M, Rau C, Jensen LJ, Bastuck S, Dumpelfeld B, Edelmann A, Heurtier MA, Hoffman V, Hoefert C, Klein K, Hudak M, Michon AM, Schelder M, Schirle M, Remor M, Rudi T, Hooper S, Bauer A, Bouwmeester T, Casari G, Drewes

- G, Neubauer G, Rick JM, Kuster B, Bork P, Russell RB, Superti-Furga G. Proteome survey reveals modularity of the yeast cell machinery. *Nature* 2006;440:631–6. [PubMed: 16429126]
17. Krogan NJ, Cagney G, Yu H, Zhong G, Guo X, Ignatchenko A, Li J, Pu S, Datta N, Tikuisis AP, Punna T, Peregrin-Alvarez JM, Shales M, Zhang X, Davey M, Robinson MD, Paccanaro A, Bray JE, Sheung A, Beattie B, Richards DP, Canadien V, Lalev A, Mena F, Wong P, Starostine A, Canete MM, Vlasblom J, Wu S, Orsi C, Collins SR, Chandran S, Haw R, Rilstone JJ, Gandhi K, Thompson NJ, Musso G, St Onge P, Ghanny S, Lam MH, Butland G, Altaf-Ul AM, Kanaya S, Shilatifard A, O'Shea E, Weissman JS, Ingles CJ, Hughes TR, Parkinson J, Gerstein M, Wodak SJ, Emili A, Greenblatt JF. Global landscape of protein complexes in the yeast *Saccharomyces cerevisiae*. *Nature* 2006;440:637–43. [PubMed: 16554755]
 18. Singh PB, Miller JR, Pearce J, Kothary R, Burton RD, Paro R, James TC, Gaunt SJ. A sequence motif found in a *Drosophila* heterochromatin protein is conserved in animals and plants. *Nucleic Acids Res* 1991;19:789–94. [PubMed: 1708124]
 19. Pray-Grant MG, Daniel JA, Schieltz D, Yates JR 3rd, Grant PA. Chd1 chromodomain links histone H3 methylation with SAGA- and SLIK-dependent acetylation. *Nature* 2005;433:434–8. [PubMed: 15647753]
 20. Marfella CG, Ohkawa Y, Coles AH, Garlick DS, Jones SN, Imbalzano AN. Mutation of the SNF2 family member Chd2 affects mouse development and survival. *J Cell Physiol* 2006;209:162–71. [PubMed: 16810678]
 21. Bouazoune K, Mitterweger A, Langst G, Imhof A, Akhtar A, Becker PB, Brehm A. The dMi-2 chromodomains are DNA binding modules important for ATP- dependent nucleosome mobilization. *EMBO J* 2002;21:2430–40. [PubMed: 12006495]
 22. Li H, Ilin S, Wang W, Duncan EM, Wysocka J, Allis CD, Patel DJ. Molecular basis for site-specific read-out of histone H3K4me3 by the BPTF PHD finger of NURF. *Nature*. 2006
 23. Shi X, Hong T, Walter KL, Ewalt M, Michishita E, Hung T, Carney D, Pena P, Lan F, Kaadige MR, Lacoste N, Cayrou C, Davrazou F, Saha A, Cairns BR, Ayer DE, Kutateladze TG, Shi Y, Cote J, Chua KF, Gozani O. ING2 PHD domain links histone H3 lysine 4 methylation to active gene repression. *Nature*. 2006
 24. Pena PV, Davrazou F, Shi X, Walter KL, Verkhusha VV, Gozani O, Zhao R, Kutateladze TG. Molecular mechanism of histone H3K4me3 recognition by plant homeodomain of ING2. *Nature*. 2006
 25. Vissers LE, van Ravenswaaij CM, Admiraal R, Hurst JA, de Vries BB, Janssen IM, van der Vliet WA, Huys EH, de Jong PJ, Hamel BC, Schoenmakers EF, Brunner HG, Veltman JA, van Kessel AG. Mutations in a new member of the chromodomain gene family cause CHARGE syndrome. *Nat Genet* 2004;36:955–7. [PubMed: 15300250]
 26. Jongmans MC, Admiraal RJ, van der Donk KP, Vissers LE, Baas AF, Kapusta L, van Hagen JM, Donnai D, de Ravel TJ, Veltman JA, Geurts van Kessel A, De Vries BB, Brunner HG, Hoefsloot LH, van Ravenswaaij CM. CHARGE syndrome: the phenotypic spectrum of mutations in the CHD7 gene. *J Med Genet* 2006;43:306–14. [PubMed: 16155193]
 27. Sakamoto I, Kishida S, Fukui A, Kishida M, Yamamoto H, Hino S, Michiue T, Takada S, Asashima M, Kikuchi A. A novel beta-catenin-binding protein inhibits beta-catenin-dependent Tcf activation and axis formation. *J Biol Chem* 2000;275:32871–8. [PubMed: 10921920]
 28. Nishiyama M, Nakayama K, Tsunematsu R, Tsukiyama T, Kikuchi A, Nakayama KI. Early embryonic death in mice lacking the beta-catenin-binding protein Duplin. *Mol Cell Biol* 2004;24:8386–94. [PubMed: 15367660]
 29. Surapureddi S, Yu S, Bu H, Hashimoto T, Yeldandi AV, Kashireddy P, Cherkaoui-Malki M, Qi C, Zhu YJ, Rao MS, Reddy JK. Identification of a transcriptionally active peroxisome proliferator-activated receptor alpha -interacting cofactor complex in rat liver and characterization of PRIC285 as a coactivator. *Proc Natl Acad Sci U S A* 2002;99:11836–41. [PubMed: 12189208]
 30. Mu J, Duan J, Makova KD, Joy DA, Huynh CQ, Branch OH, Li WH, Su XZ. Chromosome-wide SNPs reveal an ancient origin for *Plasmodium falciparum*. *Nature* 2002;418:323–6. [PubMed: 12124624]
 31. Workman JL. Nucleosome displacement in transcription. *Genes Dev* 2006;20:2009–17. [PubMed: 16882978]

32. Sims RJ 3rd, Reinberg D. Histone H3 Lys 4 methylation: caught in a bind? *Genes Dev* 2006;20:2779–86. [PubMed: 17043307]
33. Carrozza MJ, Li B, Florens L, Suganuma T, Swanson SK, Lee KK, Shia WJ, Anderson S, Yates J, Washburn MP, Workman JL. Histone H3 methylation by Set2 directs deacetylation of coding regions by Rpd3S to suppress spurious intragenic transcription. *Cell* 2005;123:581–92. [PubMed: 16286007]
34. Keogh MC, Kurdistani SK, Morris SA, Ahn SH, Podolny V, Collins SR, Schuldiner M, Chin K, Punna T, Thompson NJ, Boone C, Emili A, Weissman JS, Hughes TR, Strahl BD, Grunstein M, Greenblatt JF, Buratowski S, Krogan NJ. Cotranscriptional set2 methylation of histone H3 lysine 36 recruits a repressive Rpd3 complex. *Cell* 2005;123:593–605. [PubMed: 16286008]
35. Stokes DG, Tartof KD, Perry RP. CHD1 is concentrated in interbands and puffed regions of *Drosophila* polytene chromosomes. *Proc Natl Acad Sci U S A* 1996;93:7137–42. [PubMed: 8692958]
36. Srinivasan S, Armstrong JA, Deuring R, Dahlsveen IK, McNeill H, Tamkun JW. The *Drosophila* trithorax group protein Kismet facilitates an early step in transcriptional elongation by RNA Polymerase II. *Development* 2005;132:1623–35. [PubMed: 15728673]
37. Lutz T, Stoger R, Nieto A. CHD6 is a DNA-dependent ATPase and localizes at nuclear sites of mRNA synthesis. *FEBS Lett* 2006;580:5851–7. [PubMed: 17027977]
38. Gardner MJ, Hall N, Fung E, White O, Berriman M, Hyman RW, Carlton JM, Pain A, Nelson KE, Bowman S, Paulsen IT, James K, Eisen JA, Rutherford K, Salzberg SL, Craig A, Kyes S, Chan MS, Nene V, Shallom SJ, Suh B, Peterson J, Angiuoli S, Pertea M, Allen J, Selengut J, Haft D, Mather MW, Vaidya AB, Martin DM, Fairlamb AH, Fraunholz MJ, Roos DS, Ralph SA, McFadden GI, Cummings LM, Subramanian GM, Mungall C, Venter JC, Carucci DJ, Hoffman SL, Newbold C, Davis RW, Fraser CM, Barrell B. Genome sequence of the human malaria parasite *Plasmodium falciparum*. *Nature* 2002;419:498–511. [PubMed: 12368864]
39. LaCount DJ, Vignali M, Chettier R, Phansalkar A, Bell R, Hesselberth JR, Schoenfeld LW, Ota I, Sahasrabudhe S, Kurschner C, Fields S, Hughes RE. A protein interaction network of the malaria parasite *Plasmodium falciparum*. *Nature* 2005;438:103–7. [PubMed: 16267556]
40. Abu-Raddad LJ, Patnaik P, Kublin JG. Dual Infection with HIV and Malaria Fuels the Spread of Both Diseases in Sub-Saharan Africa. *Science* 2006;314:1603–6. [PubMed: 17158329]
41. Jacobs SA, Fischle W, Khorasanizadeh S. Assays for the determination of structure and dynamics of the interaction of the chromodomain with histone peptides. *Methods Enzymol* 2004;376:131–48. [PubMed: 14975303]
42. Otwinowski Z, Minor W. Processing of X-ray diffraction data collected in oscillation mode. *Meth. Enzymol* 1997;276:307–326.
43. Vagin A, Teplyakov A. MOLREP: an automated program for molecular replacement. *J. Appl. Cryst* 1997;30:1022–1025.
44. Perrakis A, Morris RJ, Lamzin VS. Automated protein model building with iterative structure refinement. *Nat. Struct. Biol* 1999;6:458–468. [PubMed: 10331874]
45. Emsley P, Cowtan K. COOT: Model-building tools for molecular graphics. *Acta Cryst. D* 2004;60:2126–2132. [PubMed: 15572765]
46. Murshudov GN, Vagin AA, Dodson EJ. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Cryst* 1997;D53:240–255.
47. DeLano, WL. The PyMOL molecular graphics system. DeLano Scientific; San Carlos, CA: 2004.
48. Thompson JD, Higgins DG, Gibson TJ. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* 1994;22:4673–80. [PubMed: 7984417]
49. Felsenstein J. PHYLIP (Phylogeny Inference Package) version 3.65. 2005

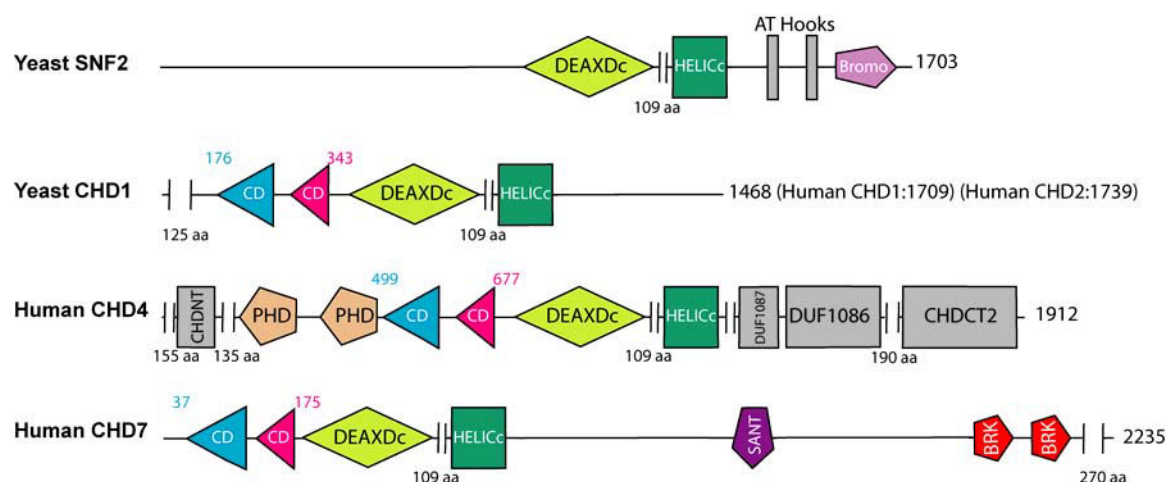


Figure 1.

Organization of the conserved domains within three distinct CHD proteins is compared with the related SNF2 protein. A conserved region with the DEAXc and HELICc domains (~450 residues) is centrally present in all these proteins, and is the ATP-dependent catalytic domain called the SNF2 domain. The residue numbers at the beginning and end of each double chromodomain region is listed above each diagram. Three distinct CHD classes exhibit variability in other conserved domains as shown for CHD1, CHD4 and CHD7. The domains are marked according to SMART nomenclature.

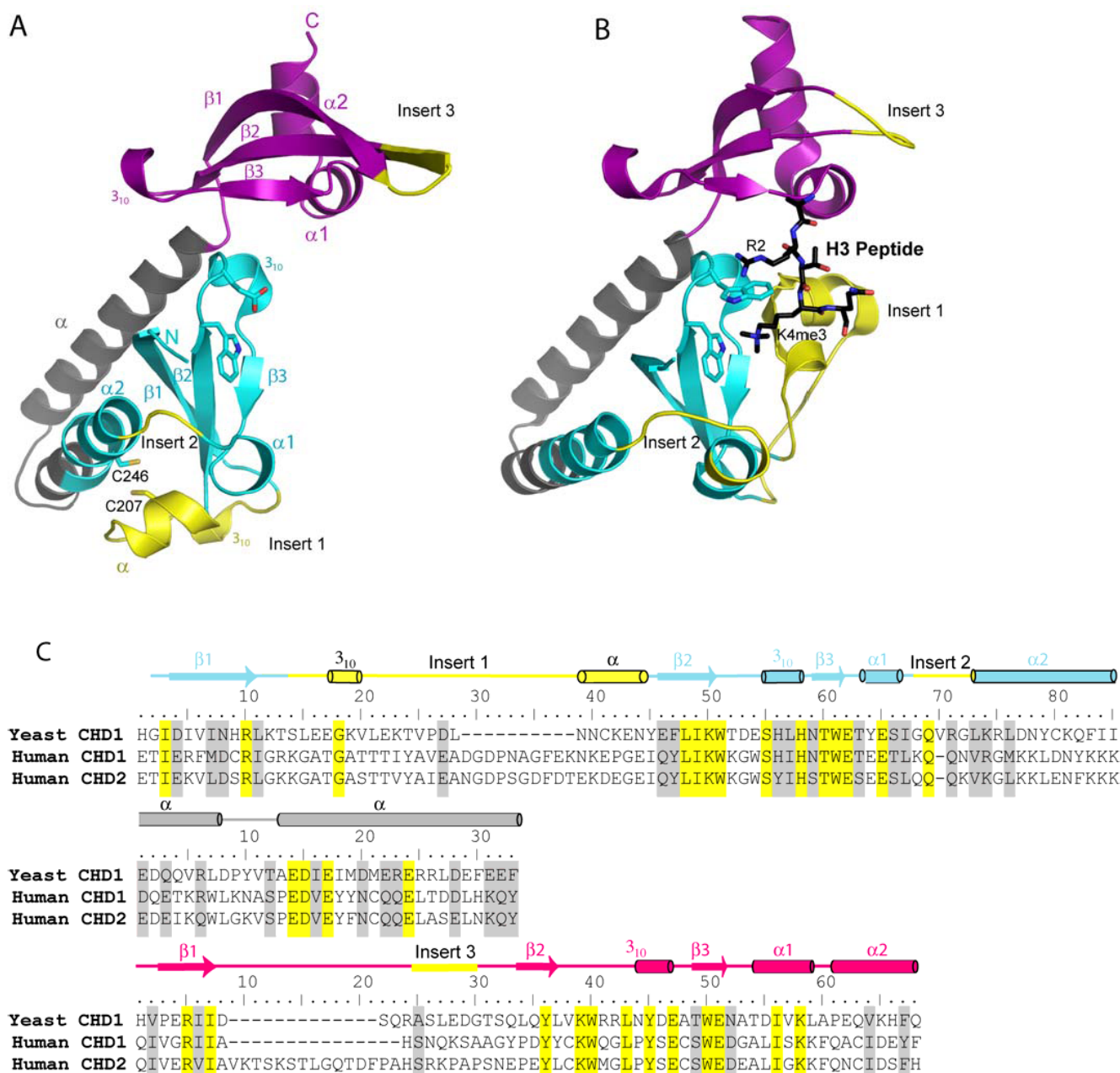
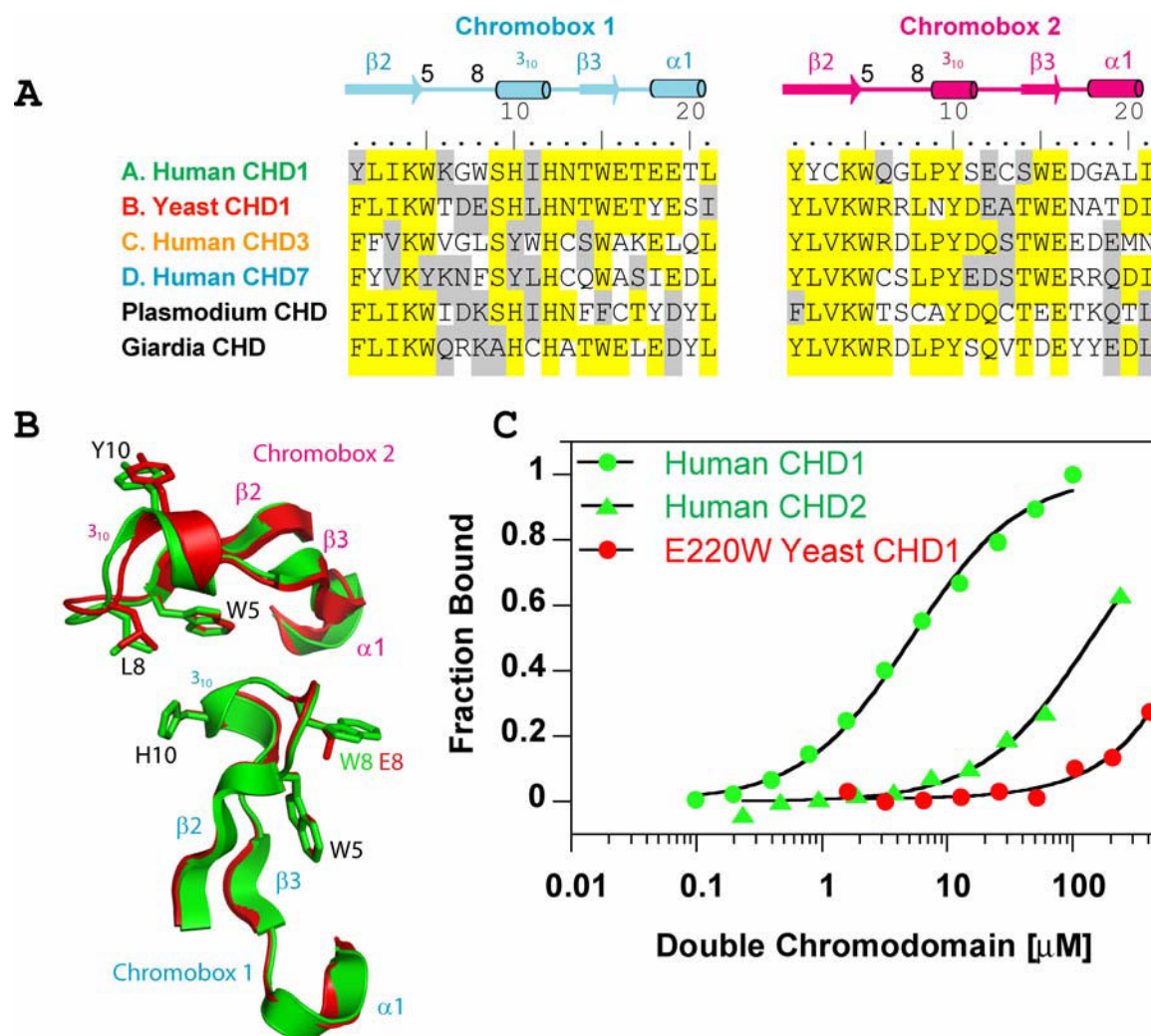


Figure 2.

The structure of yeast double chromodomains (panel A) differs from human CHD1 double chromodomain which binds the lysine 4-methylated H3 tail (panel B). Ribbon diagrams show same views. The coloring scheme is as follows, chromodomain 1 in cyan, chromodomain 2 in pink, inter-chromodomain linker in grey and the inserts relative to the HP1 or Polycomb chromodomains in yellow. The H3K4me3 peptide with sequence ARTK(me3)Q is represented by stick drawing in black. The side chains of two key residues for methyllysine binding in human CHD1 are drawn in stick representation in chromodomain 1 (Trp 322 and Trp 325 in human and Trp 217 and Glu 220 in yeast). In yeast, Cys 207 and Cys 246 may assist to attach insert 1 to $\alpha 2$ region via a disulfide bond. The aligned sequences of yeast double chromodomains and human CHD1 and CHD2 are shown for chromodomain 1 (top row), the

linker (middle row) and the chromodomain 2 (lower row). The residue numbering is arbitrary, and secondary structure diagrams are derived from human and yeast structures.

**Figure 3.**

The chromobox 1 and 2 regions fold into highly related structures. Panel A shows the sequence alignment of chromobox 1 and 2 segments of the representative CHD sequences from four classes. Panel B shows a superposition of the chromobox segments from human (green) and yeast (red) CHD1 crystal structures, showing that they arrange perpendicular to each other. The side chains of residues 5, 8 and 10 in each chromobox are shown in stick drawing. In panel C, fluorescence polarization binding assays are plotted to show human CHD1 binds specifically to the H3K4me3 peptide, whereas human CHD2 binds non-specifically and E220W mutant of yeast CHD1 does not interact with the H3K4me3 peptide.

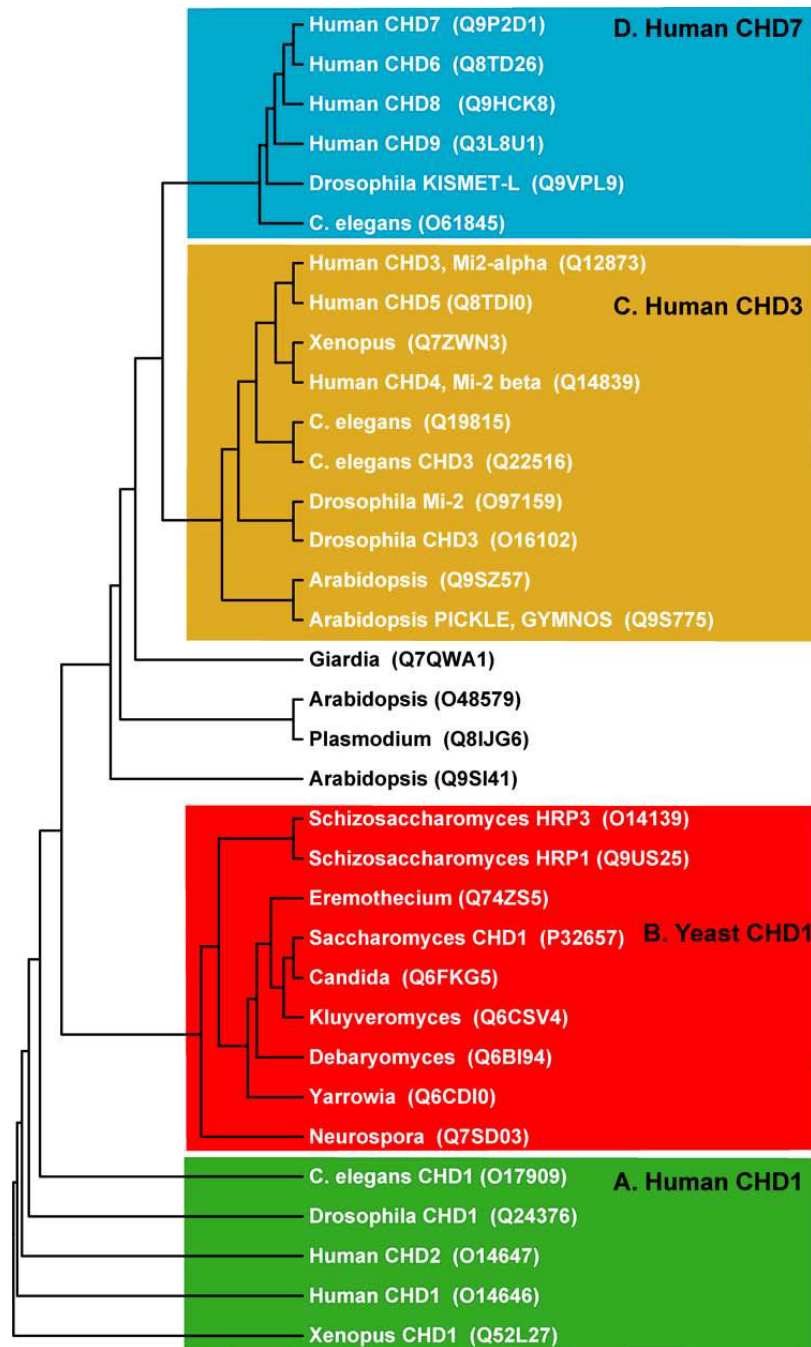


Figure 4.
A phylogenetic tree constructed using the double chromodomain regions of non-redundant CHD sequences. Four major divisions are highlighted to suggest unique double chromodomain branches. Not highlighted are the four CHD proteins which exhibit significant divergence in their sequences.

Table 1Statistics for the crystal structure of the *Saccharomyces* CHD1 double chromodomains.

Data collection	
Space group	I41
Cell dimensions	
a, b, c (Å)	136.185, 136.185, 57.516
α, β, γ (°)	90, 90, 90
Resolution (Å)	19.57(2.2)*
R_{sym} or R_{merge}	10.4(44.9)
$I/\sigma I$	10.8(1.7)
Completeness (%)	95.7(80.9)
Redundancy	3.5(2.0)
Refinement	
Resolution (Å)	2.2
No. reflections	24497
$R_{\text{work}}/R_{\text{free}}$	18.9/24.6
No. atoms	
Protein	2736
Ligand/ion	NA
Water	227
B-factors	
Protein	30.4
Ligand/ion	NA
Water	36.5
R.m.s deviations	
Bond lengths (Å)	0.023
Bond angles (°)	2.23

* Highest resolution shell is shown in parenthesis.